Perception and Iconic Memory: What Sperling Doesn't Show

IAN B. PHILLIPS

Abstract: Philosophers have lately seized upon Sperling's partial report technique and subsequent work on iconic memory in support of controversial claims about perceptual experience, in particular that phenomenology overflows cognitive access. Drawing on mounting evidence concerning postdictive perception, I offer an interpretation of Sperling's data in terms of cue-sensitive experience which fails to support any such claims. Arguments for overflow based on change-detection paradigms (e.g. Landman *et al*., 2003; Sligte *et al*., 2008) cannot be blocked in this way. However, such paradigms are fundamentally different from Sperling's and, for rather different reasons, equally fail to establish controversial claims about perceptual experience.

1. Introduction

Philosophers have lately seized upon Sperling's partial report technique and subsequent work on iconic memory in support of controversial claims about perceptual experience: that phenomenology overflows cognitive access, or that perceptual content is non-conceptual.[1] I begin by describing Sperling's experimental paradigm and the logic of arguments based upon the relevant data (§2). In doing so, I highlight a crucial and unquestioned *independence* assumption vital to any such argument, viz. that a subject's experience of the display in a partial report condition is independent of which report is cued.

Thanks to audiences in Dubrovnik, Norwich, and Budapest, especially Benj Hellie, Elizabeth Irvine, Hemdat Lerman, Fiona Macpherson, Declan Smithies, Patrick Wilken and Wayne Wu. Thanks also to Tim Bayne, Ned Block, Cecilia Heyes, Mike Martin, Hanna Pickard, Nick Shea, Henry Shelvin, Matt Soteriou, and especially to Max Coltheart for extremely helpful written comments.

Address for correspondence: Department of Philosophy, UCL, Gower Street, London, WC1E 6BT, UK.
Email: i.phillips@ucl.ac.uk

[1] The original work is reported in Sperling, 1960, and Averbach and Sperling, 1961. The term 'iconic memory', which now enjoys common currency, was introduced by Neisser (1967, p. 20) as a term for the specific 'transient visual memory' supposedly evinced by Sperling's technique. Sperling's work is used to argue for overflow by, amongst others, Block (1995, 2007, 2008), Tye (2006), Dretske (1996, pp. 151–2, fn.10; 2006), and Burge (2007). Bayne and Chalmers (2003) suggest that Sperling's data is best understood as involving a separation of access and phenomenal consciousness for complex (conjunctive) but not simple (individual letter) contents. Fodor (2007) cites it as evidence for non-conceptual content. I discuss only simple overflow arguments in what follows but my points apply to other appeals *mutatis mutandis*.

Mind & Language, Vol. 26, No. 4 September 2011, pp. 381–411.

I next review a number of psychological paradigms which cast doubt on this assumption (§3). In these paradigms the perception of an initial stimulus appears to be dramatically affected by the presence of a second stimulus—despite the fact that the second stimulus is presented a short time after initial stimulus offset and, in some cases, despite the fact that presentation occurs in a different sense modality. The temporal limits for such postdictive effects are of the same order as effective cue delays in Sperling's paradigm. Consequently, we must take seriously the possibility that initial display experience in partial report trials is subject to postdictive effects due to the subsequent cue. If so, the independence assumption is clearly false, and overflow arguments are all alike unsound.

I illustrate this (in §4) by considering the orthodox 'Stalinesque' approach to postdiction exemplified by Tye (2003). Tye himself cites two postdictive effects—'phi' apparent motion and the flash-lag illusion—in support of what he terms a 'backwards looking' account of temporal experience. Tye fails to see that, when applied to Sperling, this theory immediately makes available an interpretation of the data in terms of cue-sensitive experience. This directly undermines Tye's own use of Sperling's work to evidence phenomenal overflow. Heterodox approaches to postdiction provide scant refuge for defenders of overflow arguments.

I then ask where this discussion leaves the notion of iconic memory as defined by Sperling's paradigm (§5). I argue that Sperling at most evidences informational persistence in the sense of Coltheart (1980). I also consider and reject direct appeals to phenomenology (especially those in Block, 2007, 2008). Finally, I consider two notable sets of findings (Landman *et al*., 2003; Sligte *et al*., 2008) which Block deploys to argue for a variant of his earlier overflow claim (§6). This work plausibly avoids the difficulty just raised. However, these change-detection experiments are fundamentally different from Sperling's. As a result of these differences, they also fail to establish any controversial claims about perceptual experience.

2. Sperling's Paradigm

When presented with a stimulus display such as in Figure 1 for 15–500ms, subjects consistently report approximately 4.3 letters in their correct positions.[2] Retrospective report might seem at best a sufficient condition for experiential presentation. So we are left with a question, how many letters do subjects actually see? More generally, how rich is our perceptual world?

A familiar story relates how Sperling's pioneering work overcomes this problem and evidences the richness of experience. Sperling elicits *partial* reports from subjects by playing at random a high, medium, or low cue-tone immediately following grid presentation. The tone indicates which row is to be recalled. Once trained, subjects

[2] Sperling (1960) replicates a long attested finding here (cf. Cattell 1886). Neisser describes the 'span of apprehension' as one of the oldest 'constants' in psychology (1967, p. 41).

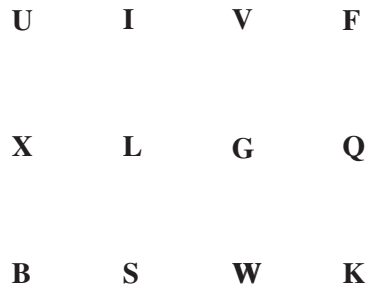

Figure 1 *Typical stimulus display used in Sperling's partial report task (Sperling, 1960; Averbach and Sperling, 1961)*

who hear a particular (e.g. high) tone consistently report approximately 3.04 letters in the corresponding (e.g. top) row.[3] This demonstrates what is known as *partial report* (PR) *superiority*: subjects are (on average) able to report more letters with respect to a row that has been cued immediately after display offset than with respect to a randomly chosen row in an uncued trial. The standard story takes something further to be demonstrated: subjects in fact *consciously perceive* at least nine letters, even if they can remember only around four in the full report (FR) condition.[4] Indeed, the standard story often simply equates this further conclusion with PR superiority. Thus, in their textbook presentation, Levine and Parkinson write, 'The increase in average number of items available, from 4.3 in whole report to 9.1 in partial report, is called the *partial report advantage*' (1994, p. 236, their emphasis).

Philosophers, led by Ned Block (1995, 2007), have been much impressed by Sperling's paradigm.[5] Dretske asks us to interpret Sperling 'as showing that subjects are perceptually conscious of more letters than they can (with such brief exposures) identify, that there is more information in their conscious experience of the letter-array than they can (in a 'full report' mode) cognitively process and report on' (2006, p. 175). Fodor (2007) deploys Sperling's work to argue for the claim that

[3] Averbach and Sperling describe their subjects as 'highly trained' (1961, pp. 196–7). Chow (1985) notes that around one hundred trials are needed to obtain a significant effect. Further training does not improve performance beyond approximately three letters.

[4] Tye (2006, p. 511) takes the conclusion to be that 'all twelve letter shapes are represented'. Similarly, Dretske (2006, p. 175) holds that '[in] the partial report condition, subjects could identify *any* letter in the entire set'. This is an exaggeration, but the precise extent of partial report superiority is not important for present purposes. All or almost all will do, as Block (1995, p. 198) is implicitly aware.

[5] Block (1995, p. 244) argues that we should think of subjects in Sperling's paradigm as being "P-conscious of all (or almost all . . .) of the letters at once, that is, jointly, and not just as blurry or vague letters, but as specific letters" but not A-conscious of 'all of them jointly, all at once'. In more recent work 'the fact of overflow is used to argue for the conclusion that the machinery of phenomenology contains more than the machinery of cognitive accessibility' (2007, p. 487). Block also draws on more recent work (see §6 below).

there is a perceptual given, i.e. that there is unconceptualized perceptual content. Finally, Tye concludes,

> My point is that at each moment, the visual experiences humans undergo are at least as rich representationally as [what Sperling describes as] sensory memories. And what Sperling's results strongly suggest is that the sensory memories are rich not only in that they represent more than their subjects actually judge to be present but also in that (typically) they represent more than their subjects are *capable* of judging to be present (2006, p. 513).

Clearly a number of assumptions are needed to move from PR superiority to conclusions about perceptual experience. Nonetheless, the central claim is as clear as it is common. As Sperling himself puts it, '*more is seen than can be remembered*' (1960, p. 1, his emphasis).

A centrally important feature in what follows is the timing of the cue-tone. Sperling investigated how soon after display offset the tone needed to be sounded to achieve PR superiority. He found that superiority rapidly decreased as delay increased (Averbach and Sperling, 1961, p. 199, Figure 5). Summarizing this finding, Averbach and Sperling report that 'only when the instruction is given within a second of the exposure do the results obtained ... differ [i.e. is PR superiority found]' (1961, p. 200). Since the results contain no data points between 300ms and 1s, this is misleading (cf. Landman *et al*. 2003, p. 150). In fact, it is widely agreed that 300ms is very close to the maximum delay which leads to partial report superiority.[6] I propose to work with this figure in what follows.

The psychological data here present an explanatory challenge, namely: how should we explain PR superiority? Sperling's idea was that there must be a kind of 'sensory memory store' encoding information about almost all the letters which endures for around 300ms, enough time for the cue to help determine which letters are transferred to a smaller capacity but more durable form of memory on the basis of which reports can be made. I do not want to question the correctness of this claim as regards sub-personal informational processing (see §5 below). However, as we have seen, Sperling's work is commonly thought to have implications regarding conscious experience. To draw conclusions about conscious experience we need to make at least two further assumptions.

The first assumption connects subjects' reports with conscious experience. The assumption is that, if a subject reports seeing the presence of some feature, then

[6] Thus, Coltheart notes that 'superiority decreased until, at cue delays greater than about 300 msec, there was no [partial report] superiority' (1980, p. 185). Likewise, Levine and Parkinson, 1994, ch. 9; Block, 2007, p. 488. In the second part of their joint paper, Averbach introduces a variation on Sperling's experiment which employs an indicator bar shown after display termination as opposed to a cue-tone. Averbach and Sperling note the similarity in method and results before offering a more precise estimate of the 'iconic storage' time (i.e. maximum delay achieving PR superiority) as 270ms (1961, Figure 17).

that is strong presumptive evidence that the subject enjoyed conscious experience as of that feature. In Sperling's experiments, 'subjects are asked to report all the letters they can see' (Gegenfurtner and Sperling, 1993, p. 845). Moreover, we have every reason to take subjects' claims that their reports are grounded in conscious experience seriously. Such subjective reports are paradigmatic criteria for awareness (see, e.g., Weiskrantz, 1997). To deny that subjective reports of this kind are probative of consciousness would be to lose experimental purchase on conscious experience. Consequently, we should grant this first assumption. Specifically, we should grant that when a subject correctly reports seeing three specific letters in a given row, they have consciously perceived all three letters in that row as the specific letters reported.

The second assumption required to justify a claim about conscious experience on the basis of PR superiority is that it is legitimate to sum partial reports to establish awareness in relevantly similar cases. In assuming this, Sperling implicitly relies on the following counterfactual:

> (CF) Any aspect of experience present in a partial report condition would have been present even if some other partial report had been cued.

Putting the assumptions together it is easy to see why Sperling's paradigm is thought to have striking consequences concerning experiential richness. By the first assumption, the letters reported in a given PR condition, say, three top row letters reported after the playing of a high tone, are consciously experienced. By (CF) these letters would also have been experienced if a different tone had been played. But if a medium tone had been played, different letters from the middle row would have been reported. These letters would also have to be regarded as aspects of the subject's experience by our first assumption. Repeating the argument with the bottom row and a low tone, we are obliged to conclude that at least nine letters are experienced despite the capacity for report in any single case being limited to half that many. In other words, (CF) legitimates summing partial reports to establish the true extent of awareness.

(CF) is held to be obvious because partial reports are elicited by a tone played just *after* stimulus offset.

> Since in [the PR condition] subjects do not know until the tone sounds which row to report on and the tone does not occur until after the array display is turned off, the fact that the subjects successfully report at least three of the four letters in the appropriate row shows that the sensory memory preserves information about the letter shapes in *all* the rows (Tye 2006, p. 511).[7]

[7] Recall that for Tye 'the visual experiences humans undergo are at least as rich representationally as [what Sperling describes as] sensory memories'. See likewise Dretske, 2006, p. 175.

In other words, (CF) is grounded in an *independence* assumption, viz. that a subject's experience of the stimulus in a PR condition is independent of which report is cued because the cue comes only *after* display offset. In what follows, I argue that experimental evidence for postdictive perception (together with subsequent reflection on the way experience is structured over the relevant timescales) provides good reason to doubt this independence assumption and with it (CF).

3. Experimental Evidence for Postdictive Perception

In a large number of experimental paradigms, the perception of an initial (target) stimulus is affected—sometimes dramatically—by a second (modulator) stimulus presented a short time after target offset. According to Choi and Scholl, these paradigms suggest,

> . . . our conscious perception of the world is not an instantaneous moment-by-moment construction, but rather is formed by integrating information presented within short temporal windows, so that new information which is obtained can influence the immediate past in our conscious awareness (2006, p. 385).

This section reviews some of these paradigms. In the next, I introduce the orthodox 'Stalinesque' interpretation, followed by two heterodox interpretations. My overarching aim is to motivate the hypothesis that, in Sperling's paradigm, the perception of the initial display is cue-sensitive, despite the fact that effective cue-tones may be delayed by up to 300ms from display offset. Consequently, I emphasise two points: first, that postdiction is well attested at timescales of 300ms; second, that auditory stimuli can act as modulators of visual targets. It should be borne in mind throughout that postdictive effects and their time-courses are extremely complex and the subject of a great deal of ongoing research. My suggestion is only that we need to take this interpretation of Sperling's data seriously, especially when it comes to assessing overflow arguments.

3.1 Backward Masking

The simplest and most extensively studied example of postdictive perception is backward visual masking. Enns and Di Lollo introduce the phenomenon as follows:

> [A] target that is highly visible when presented briefly by itself can be rendered completely invisible by the subsequent presentation of a non-target object in the same (or nearby) spatial location. 'Backward masking' of this kind has its strongest influence not when target and mask objects are presented simultaneously, as intuition might suggest, but rather when a brief temporal gap is inserted between the presentation of the target and the mask (2000, p. 345).

Figure 2 *Example of target and masking stimuli used in metacontrast masking paradigms*

Standard metacontrast masking utilises targets such as those in Figure 2 and yields a U-shaped function of visibility against stimulus onset asynchrony (SOA), with effects strongest at SOAs of around 50–100ms, decreasing sharply until effects cease to be of much significance at SOAs of above 150ms (Alpern, 1953; for reviews see Bachmann, 1994, and Breitmeyer and Öğmen, 2006).[8]

There are a vast number of different masking paradigms, the complexities of which I cannot begin to summarise (see the reviews cited above). Suffice to say that the effect is highly sensitive to a wide variety of factors. Here it is worth noting three points. Firstly, timing. The standard time course of masking is somewhat shorter than, though of much the same order as, that in Sperling's experiments. Thus, in another standard metacontrast paradigm with a target square masked by two flanking squares, Ramachandran and Cobb report, 'a characteristic ''U''-shaped function with optimum masking occurring at about 50 ms and no masking . . . at delays higher than 300 ms' (1995, p. 66). However, interestingly, when we turn to feature modulation—the modulation relevant to Sperling's task—there are an abundance of less standard, feature-masking paradigms where effects are found at significantly larger SOAs.

Particularly striking are the findings of Weisstein and Wong (1986). Weisstein and Wong were interested in the SOAs required to mask targets depending on whether the mask was perceived as figure or as ground. To investigate this they used as target a line tilted at 45° either to the left or right of the vertical; as mask they used Rubin's famous faces-vase picture which is perceived either as two faces, or as a vase, depending on which part of the figure is perceived as ground. The line was presented, its tilt randomised, for 50ms in what would be the central region of the figure (see Figure 3). Subjects were tasked to report the orientation of the

[8] Note that from the perspective of Sperling's paradigms we are primarily interested in the inter-stimulus interval (ISI), the time between display offset and cue onset. Postdictive results are often reported in terms of stimulus onset asynchrony (SOA), the time between target onset and modulator onset. To calculate the ISI we need to know target display time, since ISI = SOA − target display time. In the experiment just described a standard display time is 10ms. So the ISIs are 10ms shorter than the SOAs just given.

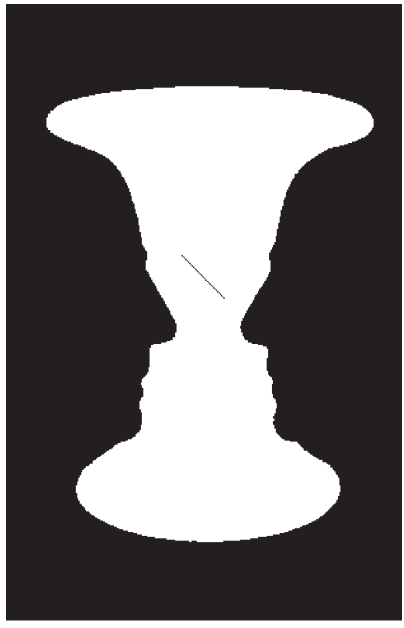

Figure 3 *Target line shown superimposed on the Rubin faces-vase mask used by Weisstein and Wong (1986). In experimental trials line and mask were onset at varying asynchronies*

line, and also whether they perceived the central region as figure or ground. The authors report their results as follows:

> When the masking area was seen as ground, the degradation of the target occurred at much longer SOAs than when the masking region was seen as a figure. The maximum masking of the figure occurred at an SOA of 300 msec. By the 700-msec SOA, the masking effect of the figure had disappeared, while the masking effect of the ground region was still strong. The maximum masking effect of ground occurred around 600 msec, some 300 msec after the maxima for the figure mask (1986, pp. 47–8).

These findings demonstrate that even with ISIs (here equal to the SOAs less 50ms) of many hundreds of milliseconds, striking postdictive effects can occur.[9]

[9] One further example. In his exploration of rod-cone interaction in the after-flash effect (where the brightness of a first flash of light is reduced by subsequent presentation of a second flash to a different retinal region), Foster (1976) reports a case of optimal masking with an SOA of 300ms (ISI = 275ms). This figure was found for an initial green flash followed by a subsequent red flash; smaller figures were found in other interactions. For discussion see Breitmeyer and Öğmen, 2006, pp. 60–1.

Secondly, cross-modality. Although masking paradigms are typically intramodal in that they involve, say, visual targets and visual modulators, cross-modal cases are also well-attested. For example, Breitmeyer investigates the effects on the 'threshold detectability of a briefly presented target stimulus consisting of a vertical sinusoidal grating' due to the presence of a masking noise. He reports that a 'two-octave-wide masking noise ... with a low spatial frequency content had its greatest masking effect on a high spatial frequency target grating when the mask followed the target by 120–180 ms' (1975, p. 297).

Thirdly, attention. Masking was once held to be entirely the result of early, autonomous visual mechanisms, as opposed to cognitive processes. However, Ramachandran and Cobb show that 'metacontrast can be strongly modulated by ''top-down'' influences such as voluntary visual attention' (1995, p. 66). That attention can be implicated in postdiction is important in relation to a postdictive interpretation of Sperling since PR superiority is very plausibly connected to the deployment of attention. For extended discussion of the role of attention in relation to iconic memory, see Phillips, 2011.

3.2 Sound-Induced Visual Bounce

Masking involves the diminution in visibility of a target. The cue in Sperling's cases, if assimilated to a postdictive modulator, does not diminish display visibility but rather helps determine *which* letters are perceived. It is important to appreciate that many postdictive effects involve such modulation rather than simple degradation. Since Sperling involves a visual display and an auditory cue, I consider an example of postdictive modulation with this structure: sound-induced visual bounce.[10]

Sekuler, Sekuler, and Lau (1997) demonstrate that roughly contemporaneous sounds can change how visual motion is seen. They do so by presenting subjects with an ambiguous display in which two dots are either seen to 'bounce' off each other and reverse direction, or alternatively to 'stream' through each other and continue on in their original directions (see Figure 4). When a sound is played at or near the point of coincidence subjects are significantly more likely to report seeing 'bouncing' as opposed to 'streaming'. Moreover, the authors report 'significant tolerance for asynchrony between sound and visual inputs: even when the sound is delayed by 150ms after coincidence, the likelihood of seeing bouncing increases' (*ibid*., p. 308).

A number of variants on this paradigm have been studied. In related work, Choi and Scholl sought to 'determine the breadth of the temporal window over which cues can affect the perception of causality' (2006, p. 388). They conclude that 'postdictive processes appear to be able to reach ''into the past'' in order to influence the construction of high-level event percepts', noting that the 'temporal windows over which such processes operate are limited to around ... 200 ms ... after the moment in question' (*ibid*., p. 393). Watanabe and Shimojo (2001) demonstrate that

[10] Thanks to Hemdat Lerman for suggesting this example.

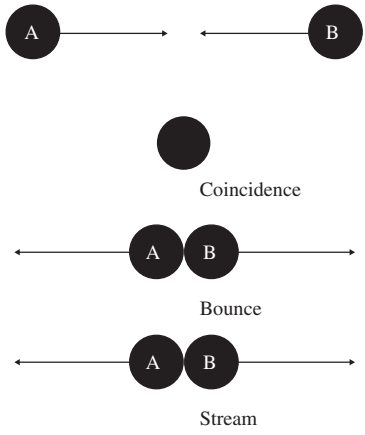

Figure 4 *Ambiguous display used by Sekuler, Sekuler, and Lau (1997) to demonstrate sound-induced visual bounce. The two dots, A and B, move towards each other and briefly coincide before moving away from each other. After coincidence, they are either seen to 'bounce' and reverse direction, or alternatively to 'stream' through each other and continue in their original directions. When a sound is played at or near the point of coincidence subjects are significantly more likely to report seeing 'bouncing' as opposed to 'streaming'*

though a sound presented 300ms after visual coincidence does not induce bounce, the 'bounce-inducing effect was attenuated when other identical sounds (auditory flankers) were presented 300ms [and perhaps up to 500ms] before and after the simultaneous sound' (p. 109). In other words, auditory context over a substantial period of time into the future (300–500ms) makes a significant difference to the effect. Interestingly, this does not seem to be a result of affecting the acoustic properties of the sound itself since these were unaffected by the presence or absence of flankers. Finally, Dufour *et al.* (2008) show that visual bounce can be induced by subliminal sounds—sounds which do not reach consciousness. This makes clear that the effect is not due to response bias but rather due to the activation of the auditory system and its involvement in visual processing. Interestingly, and in contrast to the original supraliminal experiment, in this paradigm, the effects of sounds at, or 150ms after, the moment of coincidence were not significantly different; in contrast, sounds played 150ms before coincidence had no significant effect.

3.3 Conclusions

These examples are but two of a large number of structurally similar cases where postdiction is implicated. Below I mention two other well-known examples—'colour-phi' (Kolers and von Grünau, 1976), and the flash-lag illusion (Mackay, 1958; Eagleman and Sejnowski, 2000). Further examples include the cutaneous 'rabbit' illusion (Geldard and Sherrick, 1972) as well as other kinds of apparent motion phenomena (e.g. Eagleman and Sejnowski, 2003). Many further examples

will surely be discovered. Indeed, reviewing a variety of such paradigms, Choi and Scholl 'suggest that such processing may actually be quite common in visual perception' (2006, p. 396).

What we have seen then is that the perception of an initial (target) stimulus is commonly subject to effects—sometimes dramatic ones—due to the presence of a second stimulus even though (a) the second modulator stimulus is presented several hundred milliseconds after the initial target has been displayed and offset, and (b) the modulator stimulus is in a different sense modality. In itself one might think that this provided powerful reason to doubt the independence assumption at the heart of overflow arguments based on Sperling's work. In the next section I strengthen this claim by considering the standard theoretical account of postdiction, as well as two non-standard models.

4. Postdiction and the Stream of Consciousness

This section considers three models of postdiction: (a) a backward-looking or 'Stalinesque' model which posits a delay in conscious experience; (b) an 'Orwellian' model which claims that postdictive effects result from our misremembering the nature of our initial conscious experience; and (c) an 'Extensionalist' model which arguably avoids the potential pitfalls of both Stalinesque and Orwellian approaches.

4.1 Tye's Backward-Looking Model

The simplest and orthodox way to make sense of postdictive findings is in terms of delayed perceptual consciousness. Dennett (1991, p. 115f.) calls such accounts 'Stalinesque'. This kind of account finds clear exposition in Tye's recent account of temporal experience.[11] Tye begins his discussion by considering a quite ordinary experience of two flashes or tones, *A* and *B*, perceived as occurring in close succession:

> If I have an experience of *A* followed by *B*, all in the specious present, then evidently my experience cannot be over objectively *before B* is experienced

[11] For a sophisticated formal approach to the processing naturally associated with such an account see Rao *et al.*, 2001. The authors comment, 'perception of an event is not online but rather is delayed, so that the visual system can take into account information from the immediate future before committing to an interpretation of the event' (p. 1245). See also the account of masking in Hermens *et al.*, 2008, in which the discussion begins as follows. 'Perception is not immediate. The brain processes visual information from a scene over a considerable time before a conscious percept is formed. A remarkable demonstration of this time-consuming processing comes from visual masking, in which performance on a target can be affected by a mask trailing the target by several hundred milliseconds' (p. 83). Caution, however, is needed in reading off personal-level conclusions from models of sub-personal processing.

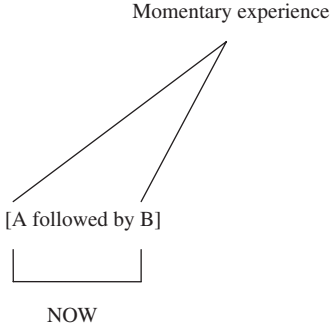

Figure 5 *A redrawn version of Tye's own illustration of his theory. The caption to the original reads: 'The experience of A followed by B in the present shown here as occurring after the specious present' (p. 89)*

> to occur, for I cannot experience *A* before *B*, unless I experience *B*. Nor can my experience begin before *A* is experienced to occur. One possibility, consistent with these claims, is that the experience of *A* followed by *B* is backward-looking. That is, it occurs with the experience of *B*, all in one go, but it represents the temporally extended period of *A*'s preceding *B*.
>
> On this account, the experience of *A* followed by *B* casts an eye backward, as it were, at what preceded it. The glance is all-in-one, however. It takes in a succession in the specious present—the time that, for the experiencer, is *now* (2003, p. 88; see Figure 5).

Tye's proposal is not the only account available which explains the intuitive data. However, it is far from idiosyncratic, being a form of the traditional specious present theory which has found many defenders over the last century or so. Indeed, Tye's diagram is almost exactly the same as C. D. Broad's (1923, p. 349). Moreover, although Tye acknowledges the availability of accounts other than his own, he thinks that we find compelling grounds for endorsing his backward-looking theory when we recognize the explanatory power it affords with respect to various psychological phenomena. Specifically, Tye cites 'colour-phi' and the flash-lag illusion. These experiments are, of course, paradigm cases of postdiction. In 'colour-phi', whether a coloured disc is seen to move (and change colour as it does so) depends on whether a nearby disc is presented within a certain temporal window.[12] Likewise,

[12] Optimal apparent motion requires an ISI of 50–60ms. However, dual or partial apparent motion, a phenomenon in which two successively presented adjacent stimuli apparently move a short distance from their actual location towards the other, occurs with ISIs of up to 200ms. Cf. the cutaneous 'rabbit' illusion which occurs with ISIs of up to 200ms (Geldard and Sherrick, 1972). Steinmann *et al.* (2000) provide a fascinating discussion of the varieties of apparent motion phenomena, and the relationship between contemporary terminology and the pioneering work of Wertheimer (1912).

in the flash-lag illusion whether a flash is perceived (accurately) as in the middle of a moving ring or (illusorily) as lagging behind it in one or other direction depends on the subsequent motion of the ring. As Eagleman and Sejnowski comment, 'the percept attributed to the time of an event is a function of what happens in the ~80ms following the event' (2000, p. 2036).

In both cases, Tye holds that the natural explanation is that, 'our brains collect information a little into the future before an experience is generated, so that what we experience as the present is in reality a little in the past' (2003, p. 91). The same story is equally applicable to the paradigms considered in the previous section. In general, we can explain postdictive perception by assuming that experience is delayed by a certain amount such that subsequent modulator stimuli can be taken into account by neural processing before the target reaches conscious awareness. The upshot is that the kind of motivational strategy that Tye invokes from postdictive findings suggests a delay in conscious experience of some several hundred milliseconds, i.e. 'NOW' in Figure 5 picks out an interval of several hundred milliseconds.[13]

We can now put things together by graphically depicting a subject's experience in Sperling's paradigm as it might look according to Tye's view of temporal consciousness. To do so we simply replace the succession: *A*, *B*; with the succession: display grid followed by cue-tone (Figure 6). On this picture, our experience of the grid is delayed until the tone has been played.[14] Given this, we no longer have any reason to assume that our experience of the grid will be independent of the tone played. After all, the tone is registered by our perceptual system *before* our experience of it and so has every opportunity to affect its character. This is precisely the possibility exploited by Tye to explain other psychological paradigms and to support his preferred model of temporal consciousness.

Above we saw that the maximum cue delay compatible with a significant effect in Sperling's task was of the order of 300ms. We also saw that this was within the range of well-attested postdictive effects, and so within the delay that Tye's theory is committed to. As a result, it is plausible to treat Sperling's paradigm in the manner above. Of course, more investigation into precise timings is required. However,

[13] Irrespective of postdiction, Tye is committed to a delay of the same order as the specious present. Tye does not venture a view as to how long he thinks the specious present lasts for. However, many specious present theorists have. Notoriously James appears to hold that 'its nucleus is probably the dozen seconds or less that have just elapsed' (1890, p. 613). Recent writers in this tradition have avoided the absurd consequences of this claim by trimming the length substantially. Thus, Dainton (2000, p. 171) tentatively offers an estimate of half a second or less. Grush recently argues for a figure of 'a few hundred milliseconds' (2007). Lockwood (2005, p. 381) thinks a second or second and a half more plausible. Kelly (2005, p. 222) declares, with implicit approval, that 'recent estimates'—he does not mentions whose—converge on a more Jamesian figure of three seconds or so. All these estimates are large enough to encompass the gap between display and tone in Sperling's paradigm.

[14] Strictly, Tye's view is only committed to the claim that there is perceptual experience of the grid which is had even once the tone has been played. But that is enough to question independence and (CF).

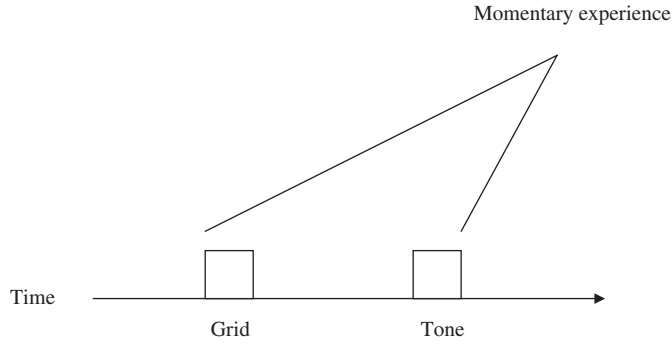

Figure 6 *Sperling's paradigm as it might look according to Tye's view of temporal consciousness*

it seems to be very much a live option on Tye's theory that we experience the grid differently depending on the tone we hear. Someone inclined to think that experience cannot outstrip cognitive access thus has a simple retort to the standard story about Sperling. We only ever experience 4–5 distinct letters as such. What explains PR superiority is that subjects experience differently depending on the tone heard. The tone does not allow subjects differentially to access a brief sensory (i.e. phenomenal) memory of a richer perceptual experience of many more letters. The tone *directly* affects which letters the subject perceives as specific letters. The fact that the tone is played after the grid is offset does not prevent this as demonstrated by the availability of a backward-looking theory of temporal awareness. Thus, the fact that the cue comes after stimulus offset does not establish (CF) or independence; and it does not legitimate the summing of partial reports.

My point is not that Sperling should somehow be assimilated to backward masking or sound induced visual-bounce. Clearly the perceptual system is under quite different constraints in each case. My point is simply that, given the coincidence of time courses, explanatory resources invoked in relation to these paradigms ought also to be considered in relation to Sperling.[15] (Someone convinced of this, and whose primary interest is in iconic memory as opposed to temporal consciousness, might wish to skip over the starred sections below, straight to section five.)

[15] That the connection between postdiction and Sperling is unremarked by philosophers interested in both is surprising given that some of the first evidence about masking arose in connection with a variant on Sperling's paradigm performed by Averbach and Coriell (1961). For brief discussion of this work and the coincidence of timing between masking and Sperling's paradigm see Levine and Parkinson (1994, pp. 237–8). Neisser's classic chapter on iconic storage comes tantalisingly close to the view proposed here. Indeed, Neisser writes that iconic memory 'underlies a number of phenomena in visual cognition, including backward masking, perceptual set, and the span of apprehension' (1967, p. 16). Unfortunately, Neisser seems ultimately to miss the view, perhaps because he does not properly distinguish sensory persistence and informational persistence, and places undue weight on introspective judgments.

4.2 Orwellian Models of Postdiction*

One might object at this point that postdiction need not be understood in terms of delayed consciousness. Two lines of thought lie behind this objection. The first cites the size of delay forced upon one in explaining postdictive effects on the Stalinesque model as a ground for doubting the model. Dennett makes this point on behalf of what he calls an 'Orwellian' interpretation of masking and apparent motion. Dennett's Orwellian argues that a 200ms delay is implausible given 'abundant evidence that responses under conscious control . . . occur with close to the minimum latencies (delays) that are physically possible' (1991, p. 122). Instead, the Orwellian proposes that we do in fact see the target stimulus even in the masking condition—we just very quickly forget it; likewise we do in fact see a stationary red disc followed by a stationary green disc in the colour-phi illusion—we just misremember what we saw as a single moving dot changing colour. The second common line of objection points to empirical evidence for just such Orwellian retention failure.

Before considering these objections, note the dialectical situation at this point in the argument. I have suggested that Sperling's data can be accounted for in terms of perceptual postdiction, a process standardly understood in Stalinesque terms. The traditional overflow interpretation of Sperling's data assumes what is, in effect, an Orwellian account (we do, in fact, see nine or more letters; we just rapidly forget most of them). My objection above to the overflow argument was, in essence, that nothing rules out giving a Stalinesque interpretation of Sperling's findings. As a result, the defender of the overflow argument must, at a minimum, establish that an Orwellian interpretation is *superior* to the Stalinesque interpretation. Simply establishing the coherence of an Orwellian account will not suffice.

However, even establishing the coherence of an Orwellian account is no mean task. An Orwellian account arguably introduces a seems/is distinction into the domain of subjective experience where many hold it does not belong. For although we can make sense of someone making a mistake about their inner life through inattention or irrationality, it is far from clear that we can make sense of experience in fact being one way and yet presenting itself to its subject in some other way. This is not the place to pursue that case.[16] Rather I want briefly to consider what reasons one might have for thinking that the Stalinesque account is untenable.

The first common objection is that conscious reaction times are incompatible with the postulated delays in perceptual consciousness. Let us grant that we can react on the basis of conscious experience of a stimulus within a period after stimulus presentation over which postdictive effects still operate. Even if this is true, the Stalinesque account has at least two responses available. Firstly, the objection assumes a crude version of the Stalinesque view on which all aspects of visual experience are subject to the same delay. Although this is true on standard versions such as Tye's,

[16] For arguments to this effect see Shoemaker, 1994. See also Martin, 2006 and discussion in Phillips, 2009.

the assumption of a constant delay is dispensable.[17] One might think that some aspects of experience, say, contour detection, were subject to a smaller delay than, say, surface completion. If a vehicle/content distinction is acceptable in this context (as Tye presumes), this need not obviously lead to any unwelcome consequences. With this additional flexibility we can allow for conscious reactions to object presence, say, even though the representation of determinate characteristics is yet to come. Secondly, the objection assumes that the kinds of delays present in laboratory conditions generalise. Since many paradigms involve substantial training, it is possible that substantial delays represent task adaptations. Rao *et al*. (2001, p. 1251) note this possibility citing support from direct input learning algorithms. I have already noted the significant training involved in iconic memory paradigms. Given this, we might well question whether such delays are present in everyday interactions.

A seemingly more radical response is to bite the bullet and grant that reactions made within the short timescales over which postdictive effects operate must be automatic responses not based on conscious experience of the relevant stimuli. It is beyond the scope of this paper properly to assess the prospects for such a view. However, it should not be dismissed out of hand. Note, in particular, that holding that all reactions at short timescales are automatic does not commit one to epiphenomenalism. The fact, if it is one, that conscious experience plays no role in the immediate generation of rapid responses to stimuli does not deprive conscious experience of an essential role in longer-term control and action-planning.

The second (at least conversationally) common objection to Stalinesque accounts points to supposed experimental evidence that postdictive paradigms, in particular visual masking, involve Orwellian retention failure. In particular, the work of Lachter and Durgin (1999) and Lachter *et al*. (2000) is often put forward as evidence for this conclusion. Lachter *et al*. studied a metacontrast paradigm with stimuli similar to those in Figure 2 above.[18] What they found was that 'the ability of subjects to distinguish such a disc/ring pair from a flickering ring is dependent [not only on SOA but] also on how soon after the stimulus they respond' (2000, p. 269). Specifically, when forced to make very rapid responses, observers were much better at detecting disc presence than when responding at leisure. The authors aver that their 'data support those who believe that these phenomena point to a kind of amnesia, a failure to remember the earlier stimuli, rather than a kind of blindness, a failure to process them in the first place' (*ibid*., p. 274).

Given such remarks, it is easy to see why Lachter *et al*. are interpreted as advocating an Orwellian account of masking. However, this last quotation is an extremely misleading summary both of the considered views of the authors and of what their work actually shows. The authors are well aware that it is hugely contestable whether the two-alternative forced-choice 'speeded' (in their case, within 480 ms)

[17] Of possible relevance here is evidence of 'flexible update' in event processing, e.g. Takei *et al*., 2008.

[18] Thanks to Tim Bayne for bringing this work to my attention.

responses they elicit should be viewed as accurate reflections of conscious awareness.[19] What their data does demonstrate is that information concerning target disc presence is *processed*. Thus, the authors do indeed evince 'a brief, but active life to *representations* of backwardly masked objects' (p. 276, my emphasis; cf. §5 below). However, the Stalinist should be quite happy with this claim. Indeed, the Stalinesque account quite explicitly invokes representations of the first stimulus, though, of course, representations that 'fail to reach consciousness'. Despite their tendentious talk of amnesia and blindness, Lachter *et al*. entirely agree with this possibility, commenting that 'the question of whether one consciously experiences these rapidly forgotten stimuli remains open' (p. 277). In particular, it is possible, they concede, that speeded responses are due to 'response routines, which can be run off ''automatically'' … [which] might be construed as ''implicit'' or ''subperceptual''' (*ibid*.). This is precisely the Stalinesque view of early responses. Their findings then carry little weight when it comes to the Orwellian/Stalinesque dispute.

In the final part of this section, I want to mention an alternative approach to postdiction which may avoid substantial delays in conscious experience. In Phillips, 2009, I develop this approach in detail as a modest realist counter-proposal to the anti-realist view found in Dennett, 1991. I mention it here primarily to illustrate how it too rejects the independence assumption required for an overflow argument based on Sperling's data. (Again, readers primarily interested in that argument may wish to skip ahead to section five.)

4.3 An Extensionalist Model of Postdiction*

Tye's account of temporal consciousness involves pulling apart the temporal structure of experience from the temporal structure of the objects presented. As Figures 5 and 6 make clear, experiential acts are momentary, yet they are said to represent unfolding successions. Though in general, experience itself need not share properties with its objects, there is a case to be made that the temporal structure of experience must match the temporal structure of its objects (see Phillips, 2009, 2010). Dainton (2000, 2008; see also Foster 1979, 1982) develops an alternative *Extensionalist* metaphysics which meets this constraint by insisting that our consciousness essentially 'extends a short distance through time' (2008, p. 631). On this model, we can endorse what Miller calls the Principle of Presentational Concurrence:

> … the time interval occupied by a content which is before the mind is the very same time interval which is occupied by the act of presenting that very content before the mind (1984, p. 107).[20]

[19] For a relevant review see Thornton and Fernandez-Duque, 2002.

[20] Miller's principle is naturally read as proposing strict, numerical identity rather than qualitative identity. But a small delay between worldly events and our experience of them seems inevitable given light transmission and minimal processing times.

Dainton's Extensionalism is intended as theory which explains our awareness of succession and change. Merely insisting that experience is extended in time does nothing to explain temporal experience for, as Husserl puts it,

> [t]he succession of sensations and the sensation of succession are not the same. Naturally we must make precisely the same objection against those who wish to trace the representation of duration and succession back to the fact of the duration and succession of psychic acts (1991, pp. 12–13).

The key claim required to make sense of temporal experience is, I suggest, not merely that experience is *extended* through time—this should be uncontroversial—but rather that there are certain durations of experience which are *metaphysically prior* to their sub-temporal-parts. In my view, this is how we should best understand the Extensionalist denial that 'our consciousness is confined to an instant' (Dainton, 2008, p. 626; cf. Soteriou, 2007).

What does it mean to say that a duration of experience is metaphysically prior to its sub-parts? It is not to deny that there are facts about instants during our stream of consciousness. It is, however, to insist that such facts are derivative. The most basic facts about our experiential lives are, in the first instance, facts about extended periods of the stream of consciousness. What is true at an instant is true only in virtue of that instant being an instant during a certain period of experience. More needs saying about metaphysical dependence, but I submit that whatever is said will sustain the considerations sketched in this section.

For present purposes, a simple analogy may help. Facts about what I am doing at some instant may depend on what I am doing over some period of time. Thus, whether I am walking or running at some instant is not fixed by a snapshot of my posture at that instant. Nor indeed is it fixed by what I do over a very brief period, e.g. $1/_{10}{}^{\text{th}}$ second, surrounding that instant. This is clear if we consider the mechanics of walking. As a textbook puts it, 'walking can be characterized as an alternating sequence of single and double support' in contrast to, say, running which 'involves alternating sequences of [single] support and nonsupport' (Enoka, 2002, p. 179). Thus, a single support phase (which is all that will be going on during certain sub-periods of periods of walking) is insufficient an occurrence on its own to determine whether someone is walking over that period. Nonetheless, someone can be walking or running at some moment in virtue of what they do over some extended period of time encompassing that instant. The metaphysically basic units of walking are significantly extended in time. The same, the Extensionalist should insist, is true of experiencing.

Dainton suggests that consciousness essentially extends as long as the specious present which he reckons to be of the order of half a second in length. If we take the metaphysically fundamental units of experience to be half a second in length, we can offer a quite different account of postdictive paradigms, and likewise Sperling's data. Consider again a standard metacontrast experiment with target and mask as in Figure 2. The orthodox view that there must be a delay in our conscious experience

to explain target invisibility is driven by the following reasoning. Consider two trials of the experiment. On the first trial only the target disc is presented; subjects report seeing it. On the second trial the same disc is presented followed by a masking ring, say, 80ms later; subjects report seeing only the masking ring. It is extremely natural to think about the situation in the first trial as follows. The stimulus is presented at time t_0; the subject then sees the disc at some later instant, $t_0 + \delta t$. On this picture we are led to ask: 'What does the subject see at $t_0 + \delta t$ in the second trial?' If we want to resist the answer being, 'The target', (as the Orwellian account claims) then we seem forced to claim that δt is a period of at least 80ms (plus further processing time)—enough time for the disc to be processed out in the light of information about the mask.

However, this natural way of thinking implicitly assumes that we can legitimately ask what is true at some instant of experience—$t_0 + \delta t$—without taking into account the nature of the subject's experience at any subsequent times. That would be legitimate if experience were analysable down to instants (as, e.g., in Tye's model). But that is precisely what the Extensionalist denies. If Extensionalism is correct, then in thinking about backward masking we have no reason to assume that what is perceived at $t_0 + \delta t$ must be the same across trials. For that instant is part of a different, metaphysically fundamental, stretch of experience. What is true over an extended period of several hundred milliseconds is that one is either presented with a target followed by a mask, or simply with a target. These are the stimuli with respect to which the basic experiential facts are determined, and upon which facts about instants are derivative. Thus, there is no reason to assume that the presence of a mask subsequent to the first stimulus is irrelevant to answering the question as to whether the first stimulus is perceived.

Certainly, it is true that if the masking ring hadn't intervened, the disc would have been seen and reported. But this has no bearing on the case where the ring was present. 'Has one seen the disc at $t_0 + \delta t$?' is not a question that one can answer independently of one's experience during the surrounding period of time. Thus, there is no reason to assume that the target will be experienced in the same way when it forms part of a different event. Our extended experience may be just of a mask, or just of a target. On this approach, whilst delay is not ruled out, it is not required.

Exactly the same approach can be taken to other postdictive effects, and, in particular, to Sperling's paradigm. On an Extensionalist view we cannot assume that prior stimulus experience is independent of the pitch of the cue-tone. The answer to the question, 'What was seen at t?', where t is the time at which the display is presented to subjects, is not answerable independently of what is presented over surrounding periods. Thus, the events 'display-followed-by-high-tone' and 'display-followed-by-low-tone' must be reckoned as different perceptual events with potentially different visual (as well as auditory) appearances. On this Extensionalist approach independence again fails. And without independence we have no reason to think that partial reports can be summed so as to testify to the whole content of any particular experience.

5. Iconic Memory

To explain his data, Sperling posited a kind of sensory memory which fades very fast following stimulus presentation but nonetheless outlives the stimulus itself. The phrase 'iconic memory' was introduced by Neisser (1967, p. 20) to refer to this specific 'transient visual memory' supposedly evinced by Sperling's technique. Subsequent writers similarly define iconic memory in terms of Sperling's paradigm. Thus, in his celebrated discussion, Coltheart tells us, '[t]he methods used in these [i.e. Sperling's] experiments define the concept ''iconic memory''' (1980, p. 185). Likewise, Levine and Parkinson note that 'partial report advantage stands as the operational definition of the icon' (1994, p. 242).

If we consider subjects' performances with respect to the grid as a whole, then unless we assume that summing partial reports is legitimate—i.e. assume (CF)—, it is not true that PR performance *is* superior. Without (CF) we simply have the same kind of limited capacity performance in both cases. All that is true is that cued row performance is, on average, better than report on a given row in a trial without a cue. Yet what we have just seen is that one can explain this finding in terms of cue-driven shifts in our perceptual experience. As a result, we need to ask how we should think of iconic memory on a postdictive account.

Coltheart notes three ways in which the phrase 'visual persistence' is used:

> This term is sometimes taken to refer to the fact that some or all of the neural components of the visual system which respond when a visual stimulus is present continue their activity for some time after the offset of the stimulus. Others use the term to refer to the fact that a visual stimulus continues to be phenomenally present, that is, 'visible,' for some time after its physical offset; and yet, again, the term is sometimes used to refer to the fact that an observer still possesses a great deal of information about the visual properties of a visual stimulus after stimulus offset (1980, p. 184).

Call these respectively: neural, phenomenological (i.e. visible), and informational persistence.

Coltheart himself equates iconic memory with informational persistence. One way of understanding this kind of persistence is as a property of representations, where these are theoretical entities in an information processing account of visual cognition. A postdictive approach to Sperling will doubtless wish to embrace this kind of persistence.[21] For the basic idea is that our experience is the result of an integration of such information across a temporal window—and the only way this integration can be made sense of in the Sperling set-up is if information of this kind concerning most of the letter identities persists until the cue arrives.

[21] Likewise, the account will want to embrace neural (and perhaps retinal) persistence. Again, such persistence establishes nothing directly about perceptual experience.

However, it cannot be emphasised enough that this informational store is at the sub-personal level and that such representations occur in many cases where no conscious awareness correlates with them.[22] Informational persistence then is certainly not the notion of iconic memory that excites Block, Tye, Dretske, and others. In this sense, we can agree with Block that 'visual representations' of almost all the items in the Sperling array are present before the cue. However, as Block puts it, 'The locus of controversy is *whether those specific representations are phenomenal*' (2007, p. 531).

One might, on the other hand, take informational persistence to be the persistence of information that has been presented in conscious experience, and so as a form of personal-level memory. Doubtless we ought to regard judgments about seen letters as grounded in conscious experience (and manifesting memory thereof). However, once we have in play the postdictive account outlined above, there is no reason to believe in such persistence regarding the specific identities of more than four or so letters. In other words, we have no good reason to move from a sub-personal informational sense of persistence to a sense which implicates rich conscious experience.

Historically, psychologists identified iconic memory with visible or phenomenological persistence.[23] A similar idea is found in the philosophical literature. Thus, Tye directly identifies iconic or sensory memory with phenomenological persistence:

> What Sperling and other psychologists call in the above case the 'visual sensory memory' is what we would call in ordinary life the look or the appearance of the array. According to Sperling, after the array has been extinguished, it *appears* still to be displayed (2006, p. 511).

Since the work of Di Lollo (1977) and Coltheart (1980), it has become increasingly accepted that visible persistence, and iconic memory as defined by Sperling's partial report task dissociate. A key piece of evidence here is that PR superiority is independent of stimulus exposure duration over a wide range (15–500ms) whereas visible persistence varies inversely with exposure duration.

It has also been suggested that iconic memory is closely related to after-images. This is difficult to believe given that subjects rarely report the presence of such after-images and, moreover, that the stimuli used are typically not bright enough to generate them.[24] Nonetheless, let us ask what relevance visible persistence or

---

[22] See again Thornton and Fernandez-Duque (2002), for example, who survey a large number of studies which converge on the view that spatiotemporally coherent mental representations can occur without awareness.

[23] See especially Neisser, 1967. For an overview see Coltheart (2009) and Coltheart and Coltheart (2010) who cite Koch (2004, p. 202) as a contemporary exponent of the visible persistence view.

[24] See respectively, Levine and Parkinson, 1994, p. 251, and Coltheart, 2009, p. 429. However, compare the discussion of the 'iconic-cue' condition in Sligte *et al.*, 2008.

after-imagery might have to our argument. In doing this, we need to bear in mind the interest in iconic memory, namely its supposed role in evincing phenomenal overflow: iconic memory is supposed to be richer in representational content than what is reportable. Even if one did think that after-images (or visibly persisting 'appearances') played some role in PR tasks, this would still leave open whether such visual phenomena have richer content than is reportable. What I have urged above is that Sperling's data provide no support for this further claim. Thus, the role of after-imagery (or visible persistence) is moot as regards overflow.

One might, of course, try and make a direct appeal to phenomenology. Indeed, Sperling himself sets out what Fodor (2007, p. 113) calls 'a pervasive phenomenological intuition', namely that 'when complex stimuli consisting of a number of letters are tachistoscopically presented, observers enigmatically insist that they have seen more than they can remember afterwards, that is; report afterwards' (1960, p. 1). Block is greatly impressed by this appeal, insisting that 'subjects exposed to a short initial stimulus have the impression that they are aware of up to 12 specific letter-shapes' (2007, p. 531). Similarly, Tye, writes that richness of experiential content (and so overflow in the relevant sense) is 'in keeping with the beliefs of the subjects about their experiences':

> Presented with a briefly displayed Sperling array, subjects believe that there are twelve *letters* in the array and indeed that they see all twelve. They also believe that the letters are all equally well-defined (2006, p. 513, emphasis in original).

However, these subjective reports are notoriously slippery. Firstly, it is implausible to think that most subjects bring to bear the necessary distinctions in making subjective reports. The crucial distinction which fails to get made at this point is between generic and specific phenomenology.[25]

Generic For every letter *c*, I saw *c* as being a letter-like form.[26]
Specific For every letter *c*, I saw *c* as being an A, or I saw *c* as being a B, or . . . I saw *c* as being a Z.

Such a distinction is no doubt borne in mind by experts. As a result, Block makes a slightly different direct phenomenological appeal, urging that 'subjects (including myself) in overflow experiments often testify that their responses are based on

[25] See Block, 2007, 2008, esp. section X. A number of commentators pick up on this distinction, esp. Byrne, Hilbert and Siegel (2007, p. 502), Papineau (2007), Sergent and Rees (2007); cf. also Spener (2007). Note, of course, that merely *seeing* an object does not require us to visually discriminate its determinate identity. We may see twelve different animals on the hillside as twelve black dots.

[26] I leave it open how generic we should think of generic phenomenology as being, i.e., as to whether subjects see 12 items as 12 letters, letter-like forms, or merely as black forms undistinguished as between letter, number, or nonsense character. For relevant empirical findings see de Gardelle *et al*., 2009.

specific phenomenology that was there all along.' (2007, p. 531; likewise 2008, p. 307; cf. Burge, 2007, p. 501). What Block fails to see, however, is that this is quite compatible with generic phenomenology relating to all unreported letters. Subjects do indeed experience the specific identities of the letters that they report. But if their experience is cue-sensitive (as I have suggested) this is quite compatible with merely generic phenomenology with respect to unreported letters. There is no reason to believe that we can sum reports. Block talks about the content being 'there all along' but in the light of the postdictive approach, this is also quite compatible with a denial that there is specific phenomenology concerning most of the displayed letters. It is also irrelevant whether or not subjects appeal to 'reading off' a visual image since there are no grounds for thinking that this image is any richer in determinate letter content than the letters that are in fact reported. So although some subjects do seem to remember in a way which involves visual imagery, there are no grounds for thinking that such imagery overflows report. In sum, neither Sperling's data nor direct appeals to phenomenology establish any controversial conclusions about phenomenal overflow. Sperling does not show that more is seen than can be remembered. What Sperling shows, as Max Coltheart put it to me, is that more is stored than can be retrieved.

6. Recent Work on Fragile Visual Short-Term Memory (VSTM)

Block (2007, 2008) cites two recent and related studies as further evidence for his views about phenomenal overflow, namely Landman *et al*., 2003, and Sligte *et al*., 2008. The cue delays involved in these experiments are significantly longer than in Sperling's partial report task: at least 1s, and with sufficient training perhaps up to 4s. Given this, it is far less plausible to account for the findings in terms of perceptual postdiction. However, even granting this, overflow arguments are not out of the woods, for these recent experiments are not simply reruns of Sperling's PR task with longer timescales. There are crucial differences. As a result of these differences, this recent work also fails to ground an overflow argument. Recall above that two assumptions were required to move from PR superiority to claims about conscious experience. The first was that, if a subject reports seeing the presence of some feature, then that is strong presumptive evidence that the subject enjoyed conscious experience as of that feature. This assumption is highly plausible in Sperling's task. As I now argue, the essential difficulty with arguments that cite Landman and Sligte's studies is that the analogous assumption is far more problematic.

The basic form of the task to be considered is as follows (see Sligte *et al*., 2008, who build on Landman *et al*., 2003, and Becker *et al*., 2000). Subjects are presented first with a 'memory-array' consisting of a number (between 4 and 32) of horizontal and vertical rectangles variously arranged. Subjects are then presented with a 'probe-array' in which all the rectangles have been rotated 90° with one possible exception. In each trial the rectangle which potentially retains its orientation is visually cued. The subjects' task is to indicate whether the cued rectangle has changed or retained

its orientation via a 'yes/no' button push. Three cue conditions are employed. First, an 'iconic-cue' condition in which the cue is presented 10ms after the initial memory-array is offset, and almost 2s before the probe-array is presented. Second, a 'retro-cue' condition in which the cue is presented 1s after the memory-array is offset, and 1s before the probe-array is presented. Finally, a 'post-cue' condition in which the probe-array is presented 900ms after the memory-array is offset, and then the cue 100ms *later* than the probe-array (see Sligte *et al*., 2008, Figure 1).

Sligte *et al*. argue that performance in the three cue conditions reveals three distinct forms of memory: '1) iconic memory with unlimited capacity, 2) a four seconds lasting fragile VSTM [visual short term memory] store with a capacity that is at least a factor of two higher than 3) the robust and capacity-limited form of VSTM' (p. 1). In the iconic-cue condition (assuming a black and white stimulus), subjects excelled and were able to report whether the rectangle had changed or not with up to 32 rectangles in the memory-array. Since this capacity was not present without a cue, the authors conclude that this evidences an iconic memory store 'with unlimited capacity'. Understood in terms of informational persistence, this inference is legitimate. Understood in terms of *phenomenal* persistence, the implicit reasoning is subject to precisely the objection above, namely that, given the possibility of postdiction, there is no reason to think of memory-array experience as independent of the iconic-cue.

Far more interesting is the novel retro-cue condition where the cue comes 1s after the memory-array has been offset (and 1s before the probe-array). Here the presence of the cue significantly improves task performance compared both to a condition where there is no cue and to the post-cue condition. Moreover, the cue comes at a delay—1s, and in other trials perhaps up to 4s, after array offset—which does not plausibly allow for an interpretation in terms of perceptual postdiction. According to Sligte *et al*., the finding, 'suggests that VSTM has an additional capacity that is however overwritten as soon as a second array (i.e. the probe array) is shown' (2008, p. 1). Landman *et al*. (2003, p. 156) propose that we should understand this capacity in terms of a long lasting form of iconic memory. Sligte *et al*. instead suggest that the proposed form of memory is fundamentally distinct from standard iconic memory as defined by Sperling's task.[27] Whatever terminology we adopt, however, the new findings can certainly be used to mount a fresh challenge to the claim that phenomenology cannot overflow access. As Sligte *et al*. put it, '[j]ust as robust VSTM forms a window on reportable and directly accessible conscious percepts, iconic memory *and fragile VSTM* could form a window on ''perception without immediate cognitive access''' (2008, p. 7, my emphasis).

In mounting this challenge, Sligte, Landman, and Block all suppose that the 'cue-advantage arises because subjects selectively transfer the cued items from iconic

[27] They give three reasons, most importantly, the significant and marked difference in capacity between iconic and retro-cue conditions just noted.

memory [or what Sligte *et al*. call 'fragile VSTM'] to . . . more durable working memory' (Landman *et al*., 2003, p. 612). In other words, they suppose that subjects must sustain a rich conscious representation of a large number of rectangles until the cue arrives so as to allow the transfer of the cued rectangle to durable working memory to facilitate comparison with the co-located rectangle in the probe-array. However, once again, to move from cue-advantage to a claim about conscious perception we need to rule out rival interpretations of the cue-advantage. In the space that remains I explain how the cue-advantage in the retro-cue condition can be accounted for without positing rich conscious experience.

To begin, note two major differences between the current paradigm and Sperling's. Firstly, Sperling's paradigm involves subjects *recalling* seen features. The current paradigm, in contrast, involves indicating whether or not a rectangle is presented with the same orientation as a previous rectangle in the same place. That is, it involves *recognising* a feature as changed/unchanged. Secondly, the Landman/Sligte paradigm is a two-alternative forced choice test. This is in marked contrast with Sperling's task where the number of letter choices is large (assuming subjects are treating all letters as available and aren't allowed blanks, there are 26 choices per square).

Given these differences, an immediate concern arises. Two-alternative forced choice tasks notoriously over-estimate conscious experience. Information which does not correlate with consciously experienced features is often picked up by the perceptual system and encoded in implicit memory (witness priming effects well-attested across a wide range of paradigms). This information can exert an impressive influence on behavioural responses in forced-choice tasks despite absence of awareness.[28] Thus, to establish an overflow interpretation of the results in Landman and Sligte's change detection task we need to establish that responses reflect explicit conscious comparison of the two arrays as opposed to a form of non-conscious priming.

To appreciate the difficulty in establishing this, consider one recent study. Using a novel recognition task with kaleidoscope displays, Voss *et al*. (2008) found remarkable dissociations between recognition accuracy and phenomenological features of explicit memory in forced-choice tests. In particular, they found that forced-choice recognition is significantly *improved* (to accuracies of over 80%) when attention is *distracted* by an independent task, and so explicit memory degraded. They conclude that there must be 'a mechanism operative for forced-choice performance that is distinct from explicit-memory processes' (2008, p. 455). In other words, their findings suggest that highly accurate responding in recognition tests does not

[28] Famously, subjects with cortical damage can learn to make reasonably reliable guesses concerning the presence of various features in their environment in certain circumstances, yet this capacity is not thought to indicate conscious awareness (Weiskrantz, 1997). Recently, Lau and Passingham (2006) demonstrated that subjective level of consciousness (whether a subject reports seeing an object) can differ in the absence of a difference in performance levels in normal subjects (what they call 'relative blindsight').

necessarily implicate explicit memory. In this light, we must take seriously the possibility that the same is true in Landman and Sligte's change detection task. If so, their results can be explained without positing rich conscious experience of the first memory-array.

The findings do, of course, reveal that rich 'visual representations' of the memory-array persist in an informational sense. As Becker *et al*. (2000) write of their closely related work on cued change detection, '[t]his finding provides evidence that the nervous system contains a more detailed (but not necessarily consciously accessible) representation of the scene than the notion of change blindness might lead one to believe' (2000, p. 283). However, as Becker *et al*. make clear, we need not think of these representations as phenomenal. Thus, as Byrne, Hilbert and Siegel comment, 'Block's appeal to Landman *et al*. must be somewhat indirect. And indeed it is, Block's argument for visible persistence is based on subjects' reports: ''[subjects say they are] continuing to maintain a visual representation of the whole array''' (2007, p. 502).

Can a direct appeal to phenomenology justify an overflow interpretation of Landman and Sligte's task? First, note that the mere fact that subjects claim to see the 'whole array' or 'all the rectangles' in the memory-array is insufficient to establish overflow as the discussion of specific versus generic phenomenology above brought out. Subjects may see all the rectangles as having some determinate orientation without it being the case that every rectangle is seen as either horizontal or seen as vertical. Second, note that *recognition* that there has been a change of orientation is compatible with a lack of conscious *recall* of the initial memory-array rectangle's orientation.[29] It is well established that people are worse at identifying items than detecting changes. This suggests that recognition of change does not always proceed by subjects consciously contrasting a presented orientation with a recalled orientation (*pace* Block, 2008, p. 305). Rather, change detection is plausibly sometimes based simply on experience of the presented probe-array: subjects may just see the rectangle's orientation as changed/unchanged. As a result subjects' claims to be consciously aware of the change (reported by Block, 2008, p. 307) need not be interpreted as evidencing their conscious experience of the relevant rectangle in the memory-array. Subjects may recognise change in the probe-array despite never having enjoyed conscious experience of the relevant rectangle in the memory-array.

Block's most persuasive phenomenological appeal is to reports that subjects who are successful in the retro-cue condition construct images to facilitate performance (2008, p. 309).[30] As Block notes, such constructed images are also implicated in temporal integration tasks (Brockmole *et al*. 2002). The problem with this appeal

[29] As Becker *et al*. put it, 'Detecting that an item at a location has changed may require far less information than is required to actually identify the original stimulus' (2000, p. 283).

[30] It is implausible to implicate image construction in Sperling's task given the short time-scales involved (Kosslyn *et al*., 2006; Brockmole *et al*., 2002; Phillips, 2011).

to imagery is that it is not clear that we need to think of the images constructed in performing these tasks as drawing on explicit memory, i.e. as arising from prior conscious experience. Instead, in the retro-cue condition, we might think of the cue as helping subjects to 'guess' at the correct orientation of rectangle to construct in their visual imagination. It is not implausible to regard trained subjects as prompting themselves to guess and thereby draw on implicit memory. This may partly explain why training improves performance in the task. It may also be what lies behind anecdotal reports that subjects in the Sligte *et al.* experiments with longer cues had to learn to 'see (and not look)', to 'relax and let it happen' (see their, 2006, cited by Block, 2008).[31] Prima facie, then, imagery-based performance in the retro-cue condition can also be accounted for in terms of implicit memory effects. Such effects can be quite striking. We should not then be surprised if accurate images can be formed by self-prompting subjects in the absence of conscious awareness of the initial memory-array.

As always, there is much more to be said.[32] My claim is only this: for all that has been said, we have no reason to deny that performance in Landman and Sligte's change detection studies results simply from sub-personal, non-conscious informational persistence. Correlatively, we have no reason to posit rich conscious awareness preserved in a fragile form of explicit memory. Despite their differences, then, neither Sperling's, nor Landman and Sligte's, experiments establish any controversial thesis about conscious awareness.

Department of Philosophy
University College, London

References

Alpern, M. 1953: Metacontrast. *Journal of the Optical Society of America*, 43, 648–57.

Averbach, E. and Coriell, A. S. 1961: Short-term memory in vision. *Bell Systems Technical Journal*, 40, 309–28.

Averbach, E. and Sperling, G. 1961: Short term storage of information in vision. In C. Cherry (ed.), *Information Theory*. Washington, DC: Butterworth & Co.

Bachmann, T. 1994: *Psychophysiology of Visual Masking: The Fine Structure of Conscious Experience*. Hauppauge, NY: Nova Science Publishers.

[31] Compare the effect of dividing attention in Voss *et al.*, 2008, discussed above.

[32] Landman and Sligte (2007) also cite neurological evidence in favour of an overflow interpretation (see also Block, 2007). Treatment of the neurological literature is beyond the scope of this article. Suffice, here, to contrast this view with that of Sergent and Rees who aver that 'strong and elaborate processing within sensory areas [of the kind Landman and Sligte point to] is not sufficient for subjective experience' (2007, p. 524).

Bayne, T. and Chalmers, D. 2003: What is the Unity of Consciousness? In A. Cleeremans (ed.), *The Unity of Consciousness*. Oxford: Oxford University Press.

Becker, M. W., Pashler, H., and Anstis, S. M. 2000: The role of iconic memory in change-detection tasks. *Perception*, 29, 273–86.

Block, N. 1995: On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, 18, 227–47.

Block, N. 2007: Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences*, 30, 481–99.

Block, N. 2008: Consciousness and cognitive access. *Proceedings of the Aristotelian Society*, 108, 289–317.

Breitmeyer, B. G. 1975: Predictions of U-shaped backward pattern masking from considerations of the spatio-temporal frequency response. *Perception*, 4, 297–304.

Breitmeyer, B. G. and Öğmen, H. 2006: *Visual Masking: Time Slices Through Conscious and Unconscious Vision*, 2nd edn. New York: Oxford University Press.

Broad, C. D. 1923: *Scientific Thought*. London: Routledge & Kegan Paul.

Brockmole, J. R., Wang, F. and Irwin, D. 2002: Temporal integration between visual images and visual percepts. *Journal of Experimental Psychology*, 28, 315–34.

Byrne, A., Hilbert, D. R. and Siegel, S. 2007: Do we see more than we can access? *Behavioral and Brain Sciences*, 30, 501–2.

Burge, T. 2007: Psychology supports independence of phenomenal consciousness. *Behavioral and Brain Sciences*, 30, 500–1.

Choi, H. and Scholl, B. J. 2006: Perceiving causality after the fact: postdiction in the temporal dynamics of causal perception. *Perception*, 35, 385–99.

Chow, S. L. 1985: Iconic store and partial report. *Memory and Cognition*, 13, 256–64.

Coltheart, M. 1980: Iconic memory and visible persistence. *Perception and Psychophysics*, 27, 183–228.

Coltheart, M. 2009: Memory, iconic. In T. Bayne, A. Cleeremans and P. Wilken (eds), *The Oxford Companion to Consciousness*. Oxford: Oxford University Press.

Coltheart, M. and Coltheart, V. 2010: Visual memories. In V. Coltheart (ed.), *Tutorials in Visual Cognition*. Hove: Psychological Press.

Dainton, B. 2000: *Stream of Consciousness*. London: Routledge and Kegan Paul.

Dainton, B. 2008: The experience of time and change. *Philosophy Compass*, 3/4, 619–38.

de Gardelle, V., Sackur, J. and Kouider, S. 2009: Perceptual illusions in brief visual presentations. *Consciousness and Cognition*, 18, 569–77.

Dennett, D. 1991: *Consciousness Explained*. Boston, MA: Little, Brown.

Di Lollo, V. 1977: Temporal characteristics of iconic memory. *Nature*, 267, 241–3.

Dretske, F. 1996: Phenomenal externalism or if meanings ain't in the head, where are qualia? *Philosophical Issues*, 7, 143–58.

Dretske, F. 2006: Perception without awareness. In T. S. Gendler and J. Hawthorne (eds), *Perceptual Experience*. Oxford: Oxford University Press.

Dufour, A., Touzalin, P., Moessinger, M., Brochard, R. and Desprès, O. 2008: Visual motion disambiguation by a subliminal sound. *Consciousness and Cognition*, 17, 790–97.

Eagleman, D. M. and Sejnowski, T. J. 2000: Motion integration and postdiction in visual awareness. *Science*, 287, 2236–8.

Eagleman, D. M. and Sejnowski, T. J. 2003: The line-motion illusion can be reversed by motion signals after the line disappears. *Perception*, 32, 963–8.

Enns, J. T. and Di Lollo, V. 2000: What's new in visual masking? *Trends in Cognitive Sciences*, 4, 345–52.

Enoka, R. M. 2002: *Neuromechanics of Human Movement*. Champaign, IL: Human Kinetics.

Fodor, J. A. 2007: The revenge of the given. In B. P. McLaughlin and J. D. Cohen (eds), *Contemporary Debates in Philosophy of Mind*. Oxford: Wiley-Blackwell.

Foster, D. H. 1976: Rod-cone interaction in the after-flash effect. *Vision Research* 16, 393–6.

Foster, J. 1979: In self-defence. In G. F. Macdonald (ed.), *Perception and Identity*. London: Macmillan.

Foster, J. 1982: *The Case for Idealism*. London: Routledge & Kegan Paul.

Gegenfurtner, K. R. and Sperling, G. 1993: Information Transfer in Iconic Memory Experiments. *Journal of Experimental Psychology* 19, 845–66.

Geldard, F. A. and Sherrick, C. E. 1972: The cutaneous 'rabbit': a perceptual illusion. *Science*, 178, 178–9.

Grush, R. 2007: Time & experience. In T. Müller (ed.) *Philosophie der Zeit: Neue analytische Ansätze*. Frankfurt: Klosterman.

Hermens, F., Luksys, G., Gerstner, W. and Herzog, M. H. 2008: Modelling spatial and temporal aspects of visual backward masking. *Psychological Review*, 115, 83–100.

James, W. 1890: *The Principles of Psychology, Vol. I*. New York: Holt.

Kelly, S. D. 2005: Temporal awareness. In D. Woodruff Smith and A. Thomasson (eds), *Phenomenology and Philosophy of Mind*. Oxford: Oxford University Press.

Koch, C. 2004: *The Quest for Consciousness*. Englewood, CO: Roberts & Company.

Kolers, P. A. and von Grünau, M. 1976: Shape and color in apparent motion. *Vision Research*, 16, 329–35.

Kosslyn, S. M., Thompson, W. L. and Ganis, G. 2006: *The Case for Mental Imagery*. Oxford: Oxford University Press.

Lachter, J. and Durgin, F. 1999: Metacontrast masking functions: a question of speed? *Journal of Experimental Psychology*, 25, 936–47.

Lachter, J., Durgin, F. and Washington, T. 2000: Disappearing percepts: evidence for retention failure in metacontrast masking. *Visual Cognition*, 7, 269–79.

Landman, R., Spekreijse, H. and Lamme, V. A. F. 2003: Large capacity storage of integrated objects before change blindness. *Vision Research*, 43, 149–64.

Landman, R. and Sligte, I. G. 2007: Can we equate iconic memory with visual awareness? *Behavioral and Brain Sciences*, 30, 512–3.

Lau, H. C. and Passingham, R. E. 2006: Relative blindsight in normal observers and the neural correlate of visual consciousness. *Proceedings of the National Academy of Sciences*, 103, 18763–8.

Levine, G. and Parkinson, S. 1994: *Experimental Methods in Psychology*. Hillsdale, NJ: Erlbaum.

Lockwood, M. 2005: *The Labyrinth of Time*. Oxford: Oxford University Press.

Mackay, D. M. 1958: Perceptual stability of a stroboscopically lit visual field containing self-luminous objects. *Nature*, 181, 507–8.

Martin, M. G. F. 2006: On being alienated. In T. S. Gendler and J. Hawthorne (eds), *Perceptual Experience*. Oxford: Oxford University Press.

Miller, I. 1984: *Husserl, Perception, and Temporal Awareness*. Cambridge, MA: MIT Press.

Neisser, U. 1967: *Cognitive Psychology*. Englewood Cliffs, NJ: Prentice-Hall.

Papineau, D. 2007: Reuniting (scene) phenomenology with (scene) access. *Behavioral and Brain Sciences*, 30, 521.

Phillips, I. B. 2009: *Experience and Time*. PhD Thesis, University College London.

Phillips, I. B. 2010: Perceiving temporal properties. *European Journal of Philosophy*, 18, 176–202.

Phillips, I. B. 2011: Attention and iconic memory. In C. Mole, D. Smithies and W. Wu (eds), *Attention: Philosophical and Psychological Essays*. Oxford: Oxford University Press.

Ramachandran, V. S. and Cobb, S. 1995: Visual attention modulates metacontrast masking. *Nature*, 373, 66–8.

Rao, R. P. N., Eagleman, D. M. and Sejnowski, T. J. 2001: Optimal smoothing in visual motion perception. *Neural Computation*, 13, 1243–53.

Sekuler, R., Sekuler, A. B. and Lau, R. 1997: Sound alters visual motion perception. *Nature*, 385, 308.

Sergent, C. and Rees, G. 2007: Conscious access overflows overt report. *Behavioral and Brain Sciences*, 30, 523–4.

Shoemaker, S. 1994: Self-knowledge and 'inner sense'. *Philosophy and Phenomenological Research*, 54, 249–314.

Sligte, I. G., Lamme, V. A. F. and Scholte, H. S. 2006: Capacity limits to awareness. Paper presented at the *Association for the Scientific Study of Consciousness*, Oxford, UK.

Sligte, I. G., Scholte, H. S. and Lamme, V. A. F. 2008: Are there multiple visual short-term memory stores? *PLoS ONE*, 3, 1–9.

Soteriou, M. 2007: Content and the stream of consciousness. *Philosophical Perspectives*, 21, 543–68.

Spener, M. 2007: Expecting phenomenology. *Behavioral and Brain Sciences*, 30, 526–7.

Sperling, G. 1960: The information available in brief visual presentations. *Psychological Monographs*, 74 (11; whole no. 498).

Steinmann, R., Pizlo, Z. and Pizlo, F. 2000: Phi is not beta, and why Wertheimer's discovery launched the Gestalt revolution. *Vision Research*, 40, 2257–64.

Takei, S., Fujisaki, W. and Nishida, S. 2008: Perceptual latency of sound-induced visual bounce [Abstract]. *Journal of Vision*, 8, 361.

Thornton, I. M. and Fernandez-Duque, D. 2002: Converging evidence for the detection of change without awareness. In J. Hyönä, D. P. Munoz, W. Heide, and R. Radach (eds), *Progress in Brain Research*, 140, 99–118.

Tye, M. 2003: *Consciousness and Persons*. Cambridge, MA: MIT Press.

Tye, M. 2006: Content, richness, and fineness of grain. In T. S. Gendler and J. Hawthorne (eds), *Perceptual Experience*. Oxford: Oxford University Press.

Voss, J., Baym, C. L. and Paller, K. A. 2008: Accurate forced-choice recognition without awareness of memory retrieval. *Learning & Memory*, 15, 454–9.

Watanabe, K. and Shimojo, S. 2001: When sound affects vision: effects of auditory grouping on visual motion perception. *Psychological Science*, 12, 109–16.

Weiskrantz, L. 1997: *Consciousness Lost and Found: A Neuropsychological Exploration*. Oxford: Oxford University Press.

Weisstein, N. and Wong, E. 1986: Figure-ground organization and the spatial and temporal responses of the visual system. In E. C. Schwab and H. C. Nusbaum (eds), *Pattern Recognition by Humans and Machines, Volume 2, Visual Perception*. Academic Press.

Wertheimer, M. 1912: Experimentelle Studien über das Sehen von Bewegung. *Zeitschrift für Psychologie*, 61, 161–265.